

Docket No.: 50195-291

D.G.
#3 7/10-02
Priority Papers
PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Application of

Takeshi ONO, et al.

Serial No.:

Group Art Unit:

Filed: March 05, 2002

Examiner:

For: VOICE RECOGNITION DEVICE



CLAIM OF PRIORITY AND
TRANSMITTAL OF CERTIFIED PRIORITY DOCUMENT

Commissioner for Patents
Washington, DC 20231

Sir:

In accordance with the provisions of 35 U.S.C. 119, Applicant hereby claims the priority of:

Japanese Patent Application Number 2001-077910, Filed March 19, 2001

cited in the Declaration of the present application. A Certified copy is submitted herewith.

Respectfully submitted,

MCDERMOTT, WILL & EMERY

Robert L. Price

Robert L. Price

Registration No. 22,685

600 13th Street, N.W.
Washington, DC 20005-3096
(202) 756-8000 RLP:kjw
Date: March 5, 2002
Facsimile: (202) 756-8087

D

50195-291
Takeshi Ono, et al
March 5, 2002

日 本 国 特 許 庁

JAPAN PATENT OFFICE

McDermott, Will & Emery

JCS79 U.S. PTO
10/087980
03/05/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office:

出 願 年 月 日
Date of Application:

2001年 3月19日

出 願 番 号
Application Number:

特願2001-077910

出 願 人
Applicant(s):

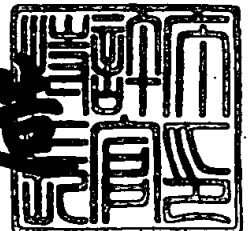
日産自動車株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年12月28日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



【書類名】 特許願

【整理番号】 NM00-00717

【あて先】 特許庁長官殿

【国際特許分類】 G10L 15/00

【発明の名称】 音声認識装置

【請求項の数】 9

【発明者】

 【住所又は居所】 神奈川県横浜市神奈川区宝町 2 番地 日産自動車株式会
社内

 【氏名】 大野 健

【発明者】

 【住所又は居所】 神奈川県横浜市神奈川区宝町 2 番地 日産自動車株式会
社内

 【氏名】 中山 沖彦

【特許出願人】

 【識別番号】 000003997

 【氏名又は名称】 日産自動車株式会社

【代理人】

 【識別番号】 100084412

 【弁理士】

 【氏名又は名称】 永井 冬紀

【手数料の表示】

 【予納台帳番号】 004732

 【納付金額】 21,000円

【提出物件の目録】

 【物件名】 明細書 1

 【物件名】 図面 1

 【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 音声認識装置

【特許請求の範囲】

【請求項 1】

音声を入力する音声入力手段と、

複数の認識対象語を記憶する記憶手段と、

前記記憶手段に記憶されている複数の認識対象語の内の、所定個数の認識対象語を表示する表示手段と、

前記表示手段に表示される認識対象語に対して前記表示手段に表示されない認識対象語よりも高い重み付けを行う重み付け手段と、

前記重み付け手段による重み付け後の認識対象語と前記入力手段により入力された音声との一致度を演算する演算手段とを備え、

一致度の演算結果に基づいて入力した音声を認識することを特徴とする音声認識装置。

【請求項 2】

請求項 1 に記載の音声認識装置において、

前記記憶手段は複数の認識対象語を所定の順に記憶しており、

前記表示手段は、前記記憶手段に前記所定の順に記憶されている複数の認識対象語の中から、前記所定個数の認識対象語を含む範囲（以下、表示範囲と呼ぶ）の認識対象語を前記所定の順に表示することを特徴とする音声認識装置。

【請求項 3】

請求項 2 に記載の音声認識装置において、

前記重み付け手段は、前記表示範囲外の認識対象語に対して、前記表示範囲から遠ざかるにしたがって前記表示範囲の認識対象語に対する重みから徐々に重みを低減することを特徴とする音声認識装置。

【請求項 4】

請求項 2 に記載の音声認識装置において、

前記表示範囲を変更する変更手段を備え、

前記重み付け手段は、前記変更手段により前記表示範囲の変更が行われた場合

のみ、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うことを特徴とする音声認識装置。

【請求項 5】

請求項 2 に記載の音声認識装置において、

前記表示範囲を変更する変更手段を備え、

前記重み付け手段は、前記変更手段により前記表示範囲の大幅な変更が行われた後に前記表示範囲の小幅な変更が行われた場合のみ、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うことを特徴とする音声認識装置。

【請求項 6】

請求項 2 に記載の音声認識装置において、

前記表示範囲を変更する変更手段を備え、

前記重み付け手段は、前記変更手段による前記表示範囲の変更にともなって前記表示範囲に現れた認識対象語の順に応じて重み付けを行い、前記表示範囲に最後に現れた認識対象語に最も高い重み付けを行うことを特徴とする音声認識装置。

【請求項 7】

請求項 2 に記載の音声認識装置において、

前記表示範囲を変更する変更手段を備え、

前記重み付け手段は、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うとともに、前記表示範囲の変更方向に存在する表示範囲外の認識対象語に対しては、前記表示範囲から遠ざかるにしたがって前記表示範囲の認識対象語に対する重みから徐々に重みを低減することを特徴とする音声認識装置。

【請求項 8】

請求項 1 ～ 7 のいずれかの項に記載の音声認識装置において、

前記演算手段により演算された一致度が高い順に所定個数の認識対象語を抽出する抽出手段を備え、

前記抽出手段により抽出された認識対象語を前記表示手段に表示して最終的な

認識対象語を選択可能にしたことを特徴とする音声認識装置。

【請求項 9】

請求項 8 に記載の音声認識装置において、

前記抽出手段により抽出された認識対象語の中に前記表示範囲外の認識対象語が含まれていない場合は、前記抽出された認識対象語の内の最も一致度の低い認識対象語と前記表示範囲外の最も一致度の高い認識対象語とを入れ替える入れ替え手段を備えることを特徴とする音声認識装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、利用者が発話した音声言葉を言葉として認識する装置に関する。

【0002】

【従来の技術】

利用者の発話する音声を認識してナビゲーション装置などの各種電子機器の入力操作を行う音声認識装置が知られている（例えば、特開 2 0 0 0 - 1 9 3 4 7 9 号公報、特開 2 0 0 0 - 2 0 3 3 5 7 号公報など参照）。

【0003】

この種の音声認識装置は、各種電子機器の入力操作に必要な言葉を予め記憶しておき（以下、これらを認識対象語と呼ぶ）、これらの認識対象語と利用者が発話した言葉とを照合し、利用者が発話した言葉と認識対象語との一致度を判定することによって、一致度が最も高い認識対象語を利用者が発話した言葉として認識している。つまり、利用者が各種電子機器の音声による入力操作に用いることができる言葉は認識対象語に限られているため、利用者は認識対象語を記憶したり、操作マニュアルを確認して発話する必要がある。

【0004】

そこで、上述した従来の音声認識装置では、利用者が音声入力操作に用いることができる認識対象語、あるいは利用者の入力が予想される認識対象語を予めモニターに表示することによって、めんどろな認識対象語を覚えたり、いちいち操作マニュアルを確認する手間を省き、利用者の使い勝手を向上させるとともに、

記憶されている認識対象語の中から利用者の発話した言葉と照合する認識対象語をモニターに表示されている認識対象語に限定することによって、認識率を向上させるという、改善がなされている。

【0005】

【発明が解決しようとする課題】

ところが、音声認識装置を例えばナビゲーション装置の音声入力操作に用いる場合には、地域名や駅名などの目的地を音声入力するための認識対象語はその数が多いので、すべてをモニターに表示することができず、一部のみを表示することになるが、利用者の発話した言葉と照合する認識対象語をモニターに表示されている認識対象語のみに限定することはできない。したがって、目的地などの数の多い認識対象語との照合により一致度を判定する場合には、認識率が低下したり、一致度の演算処理に時間がかかることになる。

【0006】

また、利用者の声の質や話し方などの個人差、雑音の有無などの周囲環境の差によって、正しい認識対象語との一致を判定することができず、誤認することがある。

【0007】

数多くの認識対象語の中から利用者の発した言葉を一致検索するときには、モニターに表示されている認識対象語も、表示されていない認識対象語も照合対象となるが、モニターに表示されている認識対象語の中から誤認が発生する場合と、表示されていない認識対象語の中から誤認が発生する場合とでは、利用者を与える影響が異なるという問題がある。利用者が発した言葉に一致する認識対象語がモニターに表示されているときに、上述した利用者の個人差や周囲環境の差などによって誤認が発生し、利用者の発した言葉を認識できないことがあると、モニターに表示されていない認識対象語に対する誤認よりも利用者により強い違和感を与えるおそれがある。

【0008】

本発明の目的は、モニターに表示されいる認識対象語に対する認識率を向上させることにある。

【 0 0 0 9 】

【課題を解決するための手段】

(1) 請求項1の発明は、音声を入力する音声入力手段と、複数の認識対象語を記憶する記憶手段と、前記記憶手段に記憶されている複数の認識対象語の内の、所定個数の認識対象語を表示する表示手段と、前記表示手段に表示される認識対象語に対して前記表示手段に表示されない認識対象語よりも高い重み付けを行う重み付け手段と、前記重み付け手段による重み付け後の認識対象語と前記入力手段により入力された音声との一致度を演算する演算手段とを備え、一致度の演算結果に基づいて入力した音声を認識する。

(2) 請求項2の音声認識装置は、前記記憶手段が複数の認識対象語を所定の順に記憶しており、前記表示手段は、前記記憶手段に前記所定の順に記憶されている複数の認識対象語の中から、前記所定個数の認識対象語を含む範囲（表示範囲）の認識対象語を前記所定の順に表示する。

(3) 請求項3の音声認識装置は、前記重み付け手段によって、前記表示範囲外の認識対象語に対して、前記表示範囲から遠ざかるにしたがって前記表示範囲の認識対象語に対する重みから徐々に重みを低減する。

(4) 請求項4の音声認識装置は、前記表示範囲を変更する変更手段を備え、前記重み付け手段によって、前記変更手段により前記表示範囲の変更が行われた場合のみ、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うようにしたものである。

(5) 請求項5の音声認識装置は、前記表示範囲を変更する変更手段を備え、前記重み付け手段によって、前記変更手段により前記表示範囲の大幅な変更が行われた後に前記表示範囲の小幅な変更が行われた場合のみ、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うようにしたものである。

(6) 請求項6の音声認識装置は、前記表示範囲を変更する変更手段を備え、前記重み付け手段によって、前記変更手段による前記表示範囲の変更にともなって前記表示範囲に現れた認識対象語の順に応じて重み付けを行い、前記表示範囲に最後に現れた認識対象語に最も高い重み付けを行うようにしたものである。

(7) 請求項7の音声認識装置は、前記表示範囲を変更する変更手段を備え、前記重み付け手段によって、前記表示範囲の認識対象語に対して前記表示範囲外の認識対象語よりも高い重み付けを行うとともに、前記表示範囲の変更方向に存在する表示範囲外の認識対象語に対しては、前記表示範囲から遠ざかるにしたがって前記表示範囲の認識対象語に対する重みから徐々に重みを低減するようにしたものである。

(8) 請求項8の音声認識装置は、前記演算手段により演算された一致度が高い順に所定個数の認識対象語を抽出する抽出手段を備え、前記抽出手段により抽出された認識対象語を前記表示手段に表示して最終的な認識対象語を選択可能にしたものである。

(9) 請求項9の音声認識装置は、前記抽出手段により抽出された認識対象語の中に前記表示範囲外の認識対象語が含まれていない場合は、前記抽出された認識対象語の内の最も一致度の低い認識対象語と前記表示範囲外の最も一致度の高い認識対象語とを入れ替える入れ替え手段を備える。

【 0 0 1 0 】

【発明の効果】

(1) 請求項1の発明によれば、表示手段に表示されている認識対象語に対する認識率を向上させることができ、利用者が発話した言葉に一致する認識対象語が表示手段に表示されているにも拘わらず、認識されないために利用者に強い違和感を与えることを避けることができる。

(2) 請求項2の発明によれば、例えば五十音順などの所定の順に表示することができ、表示手段の認識対象語の表示画面をスクロールして表示範囲を変更するときに、どちらの方向へどの程度変更すればよいかを直感でき、操作性がよくなる。

(3) 請求項3の発明によれば、利用者が表示範囲を正確に把握している必要がなく、表示範囲を変更するときに所望の認識対象語を概ね表示範囲に近づけるだけで、表示範囲に入っていなくても認識される確立が高くなり、表示範囲の設定操作に対する利用者の負担を軽減することができる。

(4) 請求項4の発明によれば、所望の認識対象語を表示手段に表示してから

発話することにより高い認識率が得られるので、面倒な操作を敢えて行う利用者を十分に満足させることができる。

(5) 請求項5の発明によれば、所望の認識対象語を確実に表示手段に表示してから発話することにより高い認識率が得られるので、面倒な操作を敢えて行う利用者を十分に満足させることができる。

(6) 利用者が表示手段の認識対象語の表示画面をスクロールして表示範囲の変更を行う場合、通常は所望の認識対象語が表示されたら変更操作を止めるので、最後に表示された認識対象語またはその近くに所望の認識対象語がある場合が多い。請求項6の発明によれば、表示範囲に現れた順に応じて重み付けを行い、表示範囲に最後に現れた語に最も高い重みを付けることによって、高い認識率が得られる。

(7) 地域名や駅名などの目的地のように認識対象語が多い場合は例えば50音順などの所定の順序に並べられていることが多く、利用者は意図する認識対象語がある方向へ認識対象語の表示画面をスクロールして表示範囲を変更する。このとき、利用者の意図する認識対象語がスクロール方向すなわち表示範囲の変更方向に存在する確立は高く、反対方向に存在する確立は低い。したがって、請求項7の発明によれば、所望の認識対象語が認識される確立を向上させることができる。

(8) 請求項8の発明によれば、表示手段に表示されている認識対象語に対する認識率をさらに向上させることができる。

(9) 請求項9の発明によれば、利用者の発話が不明瞭なために、表示範囲外に利用者の発話言葉と一致する認識対象語があるにも拘わらず、その認識対象語が重み付け後の一致度上位の認識対象語の中に含まれないという事態を避けることができ、利用者の発話が不明瞭な場合でも、最終的に選別される認識対象語の中に利用者の発話言葉に対応する認識対象語が含まれる確立が高くなり、認識率を向上させることができる。

【 0 0 1 1 】

【発明の実施の形態】

本願発明の音声認識装置をナビゲーション装置に応用した一実施の形態を説明

する。なお、本願発明はナビゲーション装置以外の多くの装置に応用することができる。

【0012】

〈発明の第1の実施の形態〉

図1に第1の実施の形態の構成を示す。ナビゲーションユニット1は、音声により入力操作を行うための音声認識ユニット2を備えており、現在地を検出するとともに目的地までの誘導経路を探索し、それらを現在地周辺の道路地図上に重畳してモニター1aに表示する。ナビゲーションユニット1には、衛星航法により現在位置を検出するGPSアンテナ3と、ナビゲーションユニット1を手動操作するためのナビリモコン4が接続される。ナビリモコン4には、モニター1aの表示内容を操作するためのジョイスティック4aと、利用者が発話の開始と終了、および発話の取り消しを指示するための発話／取り消しスイッチ4bが設けられる。また、音声認識ユニット2には、利用者が発した音声を入力するためのマイク5と、利用者に音声で応答するためのスピーカー6が接続される。

【0013】

図2は、図1に示す音声認識ユニット2の詳細を示す。音声認識ユニット2は、CPU2aやメモリ2bなどから構成される信号処理装置2cの他に、マイク5からのアナログ音声入力信号をデジタル信号に変換するA/Dコンバーター2d、デジタル音声出力信号をアナログ信号に変換するD/Aコンバーター2e、アナログ音声出力信号を増幅する出力アンプ2f、ナビゲーションユニット1とデータ通信を行うための入出力デバイス2g、認識対象語を記憶する外部記憶装置2hなどを備えている。

【0014】

図3は、第1の実施の形態の音声認識処理プログラムを示すフローチャートである。この処理プログラムは音声認識ユニット2のメモリ2bに格納される。利用者がナビリモコン4の発話／取り消しスイッチ4bを長押しすると、ナビゲーションユニット1は入力デバイス2gを介して音声認識ユニット2の信号処理装置2cへ発話開始信号を送信する。信号処理装置2cのCPU2aは、発話開始信号を受信すると図3に示す処理プログラムの実行を開始する。

【0015】

なお、この第1の実施の形態では、利用者との対話によりJRの「板橋」駅を目的地に設定する音声認識処理を例に上げて説明する。また、この例では、図4に示すような目的地として用いられる地域名や駅名などの認識対象語がナビゲーションユニット1に記憶されており、それらの認識対象語の中から利用者がナビリモコン4のジョイスティック4aを操作して、図5に示すような所定個数の認識対象語を含む任意の範囲（以下、表示範囲と呼ぶ）をモニター1aに表示する場合を説明する。

【0016】

なお、音声認識対象語は、通常、上述した目的地などの種類ごとに五十音順に記憶されており、表示範囲の認識対象語は五十音順にモニター1aに表示される。ジョイスティック4aによりモニター1aの認識対象語の表示画面をスクロールして表示範囲を変更したときは、変更後の表示範囲がナビゲーションユニット1のメモリに記憶され、次回にその種類の認識対象語を表示する場合には記憶範囲が読み出され、最初にその範囲の認識対象語がモニター1aに表示される。

【0017】

ステップ1において、現在の入力モードで用いられる認識対象語と、それらの認識対象語の内のどれがモニター1aに表示されているかをナビゲーションユニット1から入力し、認識対象語として確定する。この例では現在、目的地の入力モードが設定されているので、図4に示すような目的地として用いられる地域名や駅名などのすべての認識対象語をナビゲーションユニット1から入力するとともに、それらの認識対象語の内のモニター1aに表示されている認識対象語を入力し、認識対象語として確定する。確定した認識対象語は外部記憶装置2hに記憶する。ステップ2で、音声認識処理を開始した旨を利用者に知らせるために、外部記憶装置2hに記憶してある告知音信号をD/Aコンバーター2eおよび出力アンプ2fを介してスピーカー6へ出力し、放送する。

【0018】

ステップ3では、利用者が発話した音声の取り込みを開始する。すなわち、マイク5で入力した利用者の音声をA/Dコンバーター2dを介して信号処理装置

2 c へ入力し、外部記憶装置 2 h に収録する。信号処理装置 2 c は、発話／取り消しスイッチ 4 b が操作されるまでは、常にマイク 5 で入力し A/D コンバータ 2 d でデジタル信号に変換した雑音の平均パワーを演算している。発話／取り消しスイッチ 4 b が操作されると、それまでの平均パワーと現在の瞬間パワーとを比較し、現在の瞬間パワーが平均パワーに比べて所定値以上増加したら、利用者が発話したと判断し、音声の取り込みを開始する。

【 0 0 1 9 】

ステップ 4 で、取り込んだ音声区間部分と外部記憶装置 2 h のすべての認識対象語とを照合し、一致度の演算を開始する。一致度は、取り込んだ音声区間部分と認識対象語とがどの程度似ているかを示す値で、スコアとして得られる。この実施の形態では、スコアの値が大きいほど一致度が高いものとする。なお、ステップ 4 で一致度の演算を行っている間も、並列処理により音声の取り込みを継続する。ステップ 5 において、音声信号の瞬間パワーが所定値以下となり、その状態が所定時間続いたら利用者の発話が終了したと判断し、音声の取り込みを終了する。

【 0 0 2 0 】

ステップ 6 で一致度（スコア）の演算の終了を待って各認識対象語の一致度に重み付けを行い、一致度の高い認識対象語を抽出する。具体的には、図 4 に示すように、モニター 1 a に表示されているすべての認識対象語（表示範囲の認識対象語）に対しては、モニター 1 a に表示されていないすべての認識対象語（以下、表示範囲外の認識対象語と呼ぶ）よりも高い重みを付ける。この例では、表示範囲外の認識対象語に対しては 1. 0 の重みを付け、表示範囲の認識対象語に対しては 1. 0 よりも高い重みを付ける。そして、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位 3 個を選別し、ナビゲーションユニット 1 へ出力する。

【 0 0 2 1 】

なお、表示範囲の認識対象語に対する重みと、表示範囲外の認識対象語に対する重みは、この実施の形態の値に限定されないが、少なくとも表示範囲の認識対象語に対する重みは、表示範囲外の認識対象語に対する重みよりも大きな値とす

る。これらの重みは実験により最適な値を設定するのが望ましい。

【0022】

ナビゲーションユニット1は、図6に示すように、音声認識ユニット2から受信した一致度の高い3個の認識対象語をモニター1aに表示する。これらの認識対象語の内の「板橋」は当初、図5に示すようにモニター1aに表示されていた表示範囲の認識対象語であり、演算結果の一致度（スコア）はそれほど高くなかったが、表示範囲の認識対象語として高い重み付けをされた結果、一致度（スコア）の上位3個の中に入ったものである。

【0023】

図7は、モニター1aに表示された一致度の高い3個の認識対象語の中から、利用者が音声認識装置1との対話によって目的地を選択する手順を示す。音声認識装置1の信号処理装置2cは、一致度の最も高い認識対象語「田端（たばた）」を音声信号に変換し、D/Aコンバーター2eおよび出力アンプ2fを介してスピーカー6から放送する。この放送を聞いた利用者は、「田端」は目的地に設定しようとしている駅名ではないので、ナビリモコン4の発話／取り消しスイッチ4bの短押しを行う。ナビゲーションユニット1は発話／取り消しスイッチ4bの短押し操作を検出し、音声認識ユニット2へ取り消し信号を送信する。

【0024】

取り消し信号を受信した音声認識ユニット2の信号処理装置2cは、一致度（スコア）が第2順位の「板橋（いたばし）」を音声信号に変換し、D/Aコンバーター2eおよび出力アンプ2fを介してスピーカー6から放送する。この放送を聞いた利用者は、「板橋」は目的地に設定しようとしている駅名であり、「目的地設定」という発話を行い、目的地を確定する。音声認識ユニット2の信号処理装置2cは利用者からの「目的地設定」の音声入力を認識し、「板橋」を利用者が確定した目的地としてナビゲーションユニット1へ送信する。ナビゲーションユニット1は「板橋」を目的地として設定する。

【0025】

このように、利用者によってモニター1aに表示された認識対象語、すなわち表示範囲の認識対象語に含まれる言葉「板橋」を発話したときに、この「板橋」

が一致度（スコア）の高い上位 3 個の認識対象語に含まれる確立が高くなり、発話した言葉が表示範囲の認識対象語の一つとしてモニター 1 a に表示されているにも拘わらず、発話した言葉が認識されずに利用者に違和感を与えることが避けられる。

【 0 0 2 6 】

＜発明の第 2 の実施の形態＞

各認識対象語の一致度に対する他の重み付け方向を説明する。なお、この第 2 の実施の形態の構成は図 1 および図 2 に示す第 1 の実施の形態の構成と同様であり、説明を省略する。また、この第 2 の実施の形態の動作についても、重み付け方法以外は図 3 に示す第 1 の実施の形態の動作を同様であり、説明を省略する。

【 0 0 2 7 】

上述した第 1 の実施の形態では、図 4 に示すように、表示範囲のすべての認識対象語「軍畑」～「稻城長沼」に対しては 1. 0 よりも高い所定の重みを付け、表示範囲外のすべての認識対象語に対しては 1. 0 の重みを付ける方法を説明した。この第 2 の実施の形態では、図 8 に示すように、表示範囲のすべての認識対象語に対しては 1. 0 よりも高い所定の重みを付ける。一方、表示範囲外の認識対象語に対しては、図 8 に示すように、表示範囲から遠ざかるにしたがって表示範囲の認識対象語の重みから徐々に重みを低減し、最終的に 1. 0 の重みとする。

【 0 0 2 8 】

これにより、利用者が表示範囲を正確に把握している必要がなく、例えば図 5 に示すような目的地選択画面において利用者がジョイスティック 4 a により表示範囲を変更できる場合に、所望の認識対象語を概ね表示範囲に近づけるだけで、表示範囲に入っていなくても認識される確立が高くなり、所望の認識対象語を確実にモニター 1 a に表示しなければ認識確立を上げることができない上述した第 1 の実施の形態に比べ、表示範囲の設定操作に対する利用者の負担を軽減することができる。

【 0 0 2 9 】

＜発明の第 3 の実施の形態＞

上述した第1の実施の形態では、表示範囲の認識対象語に対して表示範囲外の認識対象語に対する重みよりも高い重みを付け、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位3個を選別して表示する例を示した。この第3の実施の形態では、最終的に選別された一致度上位3個の認識対象語の中に、表示範囲外の認識対象語が含まれていない場合は、重み付け後の表示範囲外の認識対象語の中から最も一致度（スコア）が高い語を抽出し、表示範囲の一致度上位3個の認識対象語の内の最も一致度（スコア）の低い語と入れ替える。

【0030】

これにより、利用者の発話が不明瞭なために、表示範囲外に利用者の発話言葉と一致する認識対象語があるにも拘わらず、その認識対象語が重み付け後の一致度上位3個の認識対象語の中に含まれないという事態を避けることができ、利用者の発話が不明瞭な場合でも、最終的に選別される認識対象語の中に利用者の発話言葉に対応する認識対象語が含まれる確立が高くなり、認識率を向上させることができる。

【0031】

さらに具体的には、図5に示す目的地選択画面において、利用者が「あっ、亀有」というような不要語を含む言葉を発話した場合に、発話語が「あ」から始まっているために最終的な選別結果に「亀有」が含まれないことがあるが、この第3の実施の形態の選別方法によれば、最終的な選別結果に「亀有」が含まれる確立が高くなる。

【0032】

なお、この第3の実施の形態の構成は図1および図2に示す第1の実施の形態の構成と同様であり、説明を省略する。また、この第3の実施の形態の動作は、重み付け後の一致度上位3個の選別方法を除き、図3に示す第1の実施の形態の動作と同様であり、説明を省略する。

【0033】

《発明の第4の実施の形態》

上述した第1の実施の形態では、表示範囲に含まれるか否かによって認識対象

語に異なる重み付けを行う例を示したが、この第4の実施の形態では、認識対象語に対する重み付けを行わない。

【0034】

この第4の実施の形態では、認識対象語に重み付けを行わず、利用者の発話言葉と照合して演算した一致度（スコア）が上位3個の認識対象語を選別し、その選別した認識対象語の中に表示範囲の内と外の両方の認識対象語が含まれているかどうかを確認する。一致度上位3個の認識対象語の中に表示範囲と表示範囲外の両方の認識対象語が混在しておらず、表示範囲内または表示範囲外のいずれか一方の認識対象語のみが含まれている場合は、選別結果に含まれていない方の認識対象語の中から一致度（スコア）が最も高い語を抽出し、選別結果の認識対象語の中の最も一致度（スコア）の低い語と入れ替える。

【0035】

これにより、利用者が発話した言葉に対応する認識対象語が表示されているにも拘わらず、その認識対象語が最終的に選別されず、利用者に強い違和感を与えるような事態を避けることができ、利用者の発話が不明瞭な場合でも、最終的に選別される認識対象語の中に利用者の発話言葉に対応する認識対象語が含まれる確立が高くなり、認識率を向上させることができる。

【0036】

なお、この第4の実施の形態の構成は図1および図2に示す第1の実施の形態の構成と同様であり、説明を省略する。また、この第3の実施の形態の動作は、一致度上位3個の認識対象語の選別方法を除き、図3に示す第1の実施の形態の動作と同様であり、説明を省略する。

【0037】

《発明の第5の実施の形態》

利用者が音声認識による操作を行う前に表示範囲の変更を行った場合のみ、表示範囲の認識対象語に対して表示範囲外の認識対象語に対する重みよりも高い重みを設定する、第5の実施の形態を説明する。なお、この第5の実施の形態の構成は図1および図2に示す第1の実施の形態の構成と同様であり、説明を省略する。

【 0 0 3 8 】

図 9 は、第 5 の実施の形態の音声認識処理プログラムを示すフローチャートである。なお、図 3 に示す第 1 の実施の形態の音声認識処理と同様な処理を行うステップに対しては同一のステップ番号を付して相違点を中心に説明する。

【 0 0 3 9 】

音声認識処理開始後のステップ 1 1 において、ナビゲーションユニット 1 の操作履歴を取得する。このとき、発話／取り消しスイッチ 4 b の操作から遡って所定数の操作の履歴を取得する。ジョイスティック 4 a により認識対象語の表示範囲が変更されていた場合には、操作（ジョイスティック操作）とその結果（表示範囲変更）の両方が履歴として残されている。信号処理装置 2 c はそれらの操作履歴をメモリ 2 b に記憶する。

【 0 0 4 0 】

一致度演算終了後のステップ 6 A において、まずメモリ 2 b に記憶されている操作履歴から発話／取り消しスイッチ 4 b を操作する前の操作内容を確認する。発話／取り消しスイッチ 4 b の操作前に表示範囲の変更操作が行われた記録がある場合は、第 1 の実施の形態で上述したように、表示範囲の認識対象語に対しては 1. 0 よりも高い所定の重みを付け、表示範囲外の認識対象語に対しては 1. 0 の重みを付ける。そして、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位 3 個を選別する。一方、発話／取り消しスイッチ 4 b の操作前に表示範囲の変更操作が行われていない場合は、認識対象語に対する重み付けを行わず、表示範囲と表示範囲外のすべての認識対象語の中から一致度が高い上位 3 個の認識対象語を選別する。

【 0 0 4 1 】

この第 5 の実施の形態によれば、利用者が所望の認識対象語をモニター 1 a に表示してから発話した場合のみ、表示範囲の認識対象語に表示範囲外の認識対象語よりも高い重みを付け、表示範囲の認識対象語に対する認識率を向上させる。しかし、利用者が所望の認識対象語をモニター 1 a に表示する操作を行わずに発話した場合は、表示範囲の認識対象語に対する重み付けを行わず、表示範囲と表示範囲外のすべての認識対象語を平等に取り扱う。これにより、所望の認識対象

語をモニター 1 a に表示してから発話するので高い認識率が得られ、面倒な操作を敢えて行う利用者を十分に満足させることができる。

【0042】

《発明の第 6 の実施の形態》

表示範囲の大幅な変更を行った後に表示範囲の小幅な変更を行い、その後に音声認識による操作を行う場合のみ、表示範囲の認識対象語に対して表示範囲外の認識対象語に対する重みよりも高い重みを設定する、第 6 の実施の形態を説明する。なお、この第 6 の実施の形態の構成は図 1 および図 2 に示す第 1 の実施の形態の構成と同様であり、説明を省略する。

【0043】

この第 6 の実施の形態の動作は、図 9 に示す第 5 の実施の形態の動作の内のステップ 6 A の一部が異なるのみであり、図示を省略する。一致度演算終了後のステップ 6 A において、まずメモリ 2 b に記憶されている操作履歴から発話／取り消しスイッチ 4 b を操作する前の操作内容を確認する。発話／取り消しスイッチ 4 b の操作前に、表示範囲の大幅な変更操作が行われ、その後に表示範囲の小幅な変更操作が行われた記録がある場合は、第 1 の実施の形態で上述したように、表示範囲の認識対象語に対しては 1. 0 よりも高い所定の重みを付け、表示範囲外の認識対象語に対しては 1. 0 の重みを付ける。そして、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位 3 個を選別する。一方、発話／取り消しスイッチ 4 b の操作前に表示範囲の変更操作が行われていない場合は、認識対象語に対する重み付けを行わず、表示範囲と表示範囲外のすべての認識対象語の中から一致度が高い上位 3 個の認識対象語を選別する。

【0044】

この第 6 の実施の形態では、1 回の表示範囲の変更操作において、例えば、表示範囲の認識対象語がすべて入れ替わる変更を表示範囲の大幅な変更とし、表示範囲の一部の認識対象語が入れ替わる変更を表示範囲の小幅な変更とする。

【0045】

この第 6 の実施の形態によれば、所望の認識対象語をモニター 1 a に確実に表

示してから発話するという、面倒な操作を敢えて行う利用者に報いるだけの充分に高い認識率が得られる。

【0046】

〈発明の第7の実施の形態〉

利用者が音声認識による操作を行う前に、表示範囲の変更を行った場合のみ、表示範囲の認識対象語に対して表示範囲外の認識対象語に対する重みよりも高い重みを、表示範囲の変更方向に応じて設定するようにした第7の実施の形態を説明する。なお、この第7の実施の形態の構成は図1および図2に示す第1の実施の形態の構成と同様であり、説明を省略する。

【0047】

この第7の実施の形態の動作は、図9に示す第5の実施の形態の動作の内のステップ6Aの一部が異なるのみであり、図示を省略する。一致度演算終了後のステップ6Aにおいて、まずメモリ2bに記憶されている操作履歴から発話／取り消しスイッチ4bを操作する前の操作内容を確認する。発話／取り消しスイッチ4bの操作前に表示範囲の変更操作が行われた記録がある場合は、図10に示すように、表示範囲の認識対象語に対して1.0よりも高い重みを、ジョイスティック4aによる表示範囲のスクロール方向、すなわち表示範囲の変更方向にある認識対象語ほど高い重みを付け、表示範囲外の認識対象語に対しては1.0の重みを付ける。そして、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位3個を選別する。一方、発話／取り消しスイッチ4bの操作前に表示範囲の変更操作が行われていない場合は、認識対象語に対する重み付けを行わず、表示範囲と表示範囲外のすべての認識対象語の中から一致度が高い上位3個の認識対象語を選別する。

【0048】

例えば図11に示すような地域名や駅名などの目的地の表示画面において、ジョイスティック4aにより目的地の表示範囲をスクロールする場合、所望の認識対象語が現れたらスクロールを止めることから、最後に現れた目的地、つまり認識対象語か、もしくはその近傍の認識対象語が利用者の所望の認識対象語である。この第7の実施の形態では、表示範囲のすべての認識対象語の重みを一律に設

定せず、スクロール方向すなわち表示範囲の変更方向にある認識対象語ほど高い重みを設定する。換言すれば、表示範囲に現れた順に応じて重み付けを行い、最後に表示範囲に現れた認識対象語に対して最も高い重みを設定する。図 1 1 に示す表示範囲の例では、利用者の所望の目的地（認識対象語）「市ヶ谷」は、スクロールを止めたときに最後に現れた目的地の近くにあり、高い重みが設定されるため、認識結果として現れる確立を高くすることができる。

【 0 0 4 9 】

《発明の第 8 の実施の形態》

利用者が音声認識による操作を行う前に、表示範囲の変更を行った場合のみ、表示範囲の認識対象語に対して表示範囲外の認識対象語に対する重みよりも高い重みを設定するとともに、表示範囲外の認識対象語の内の表示範囲の変更方向にある認識対象語に対しては、表示範囲から遠ざかるにしたがって表示範囲の認識対象語の重みから徐々に重みを低減するようにした第 8 の実施の形態を説明する。なお、この第 8 の実施の形態の構成は図 1 および図 2 に示す第 1 の実施の形態の構成と同様であり、説明を省略する。

【 0 0 5 0 】

この第 8 の実施の形態の動作は、図 9 に示す第 5 の実施の形態の動作の内のステップ 6 A の一部が異なるのみであり、図示を省略する。一致度演算終了後のステップ 6 A において、まずメモリ 2 b に記憶されている操作履歴から発話／取り消しスイッチ 4 b を操作する前の操作内容を確認する。発話／取り消しスイッチ 4 b の操作前に表示範囲の変更操作が行われた記録がある場合は、図 1 2 に示すように、表示範囲の目的地（認識対象語）に対して 1. 0 よりも高い所定の重みを設定するとともに、表示範囲外の目的地（認識対象語）の内のジョイスティック 4 a による表示範囲のスクロール方向、すなわち表示範囲の変更方向にある目的地（認識対象語）に対しては、表示範囲から遠ざかるにしたがって表示範囲の目的地（認識対象語）の重みから徐々に重みを低減し、最終的に 1. 0 の重みに収束させる。そして、各認識対象語の一致度（スコア）に重みを乗じ、重み付け後の認識対象語の中から一致度（スコア）の高い上位 3 個を選別する。一方、発話／取り消しスイッチ 4 b の操作前に表示範囲の変更操作が行われていない場合

は、認識対象語に対する重み付けを行わず、表示範囲と表示範囲外のすべての認識対象語の中から一致度が高い上位 3 個の認識対象語を選別する。

【 0 0 5 1 】

地域名や駅名などの目的地のように認識対象語が多い場合は 5 0 音順に並べられていることが多く、利用者はジョイスティックや方向キーにより利用者の意図する認識対象語がある方向へ画面をスクロールする。このとき、利用者の意図する認識対象語がスクロール方向に存在する確立は高く、反対方向に存在する確立は低い。そこで、この第 8 の実施の形態では、表示範囲のスクロール方向、つまり表示範囲の変更方向にある表示範囲外の認識対象語に対しては、表示範囲から遠ざかるにしたがって表示範囲の認識対象語の重みから徐々に重みを低減し、最終的に 1. 0 の重みに収束させる。一方、スクロール操作により表示範囲を通り過ぎてしまった認識対象語、つまり表示範囲の変更方向と反対の方向の表示範囲外の認識対象語に対しては、利用者の意図する認識対象語である可能性が低い語と見なし、1. 0 の重みを付ける。これにより、所望の認識対象語が認識される確立を向上させることができる。

【 0 0 5 2 】

以上の実施の形態の構成において、マイク 5 が音声入力手段を、外部記憶装置 2 h が記憶手段を、モニター 1 a が表示手段を、音声入力ユニット 2 の CPU 2 a が重み付け手段、演算手段、変更手段、抽出手段および入れ替え手段をそれぞれ構成する。

【図面の簡単な説明】

【図 1】 第 1 の実施の形態の構成を示す図である。

【図 2】 図 1 に示す音声認識ユニットの詳細を示す図である。

【図 3】 第 1 の実施の形態の音声認識処理プログラムを示すフローチャートである。

【図 4】 第 1 の実施の形態の認識対象語の重み付け方法を示す図である。

【図 5】 表示範囲の認識対象語の表示例を示す図である。

【図 6】 重み付け後の一致度上位 3 個の認識対象語の表示例を示す図である。

【図 7】 一致度の高い認識対象語からいずれかを選択する手順を示す図である。

【図 8】 第 2 の実施の形態の認識対象語の重み付け方法を示す図である。

【図 9】 第 5 の実施の形態の音声認識処理プログラムを示すフローチャートである。

【図 1 0】 第 7 の実施の形態の認識対象語の重み付け方法を示す図である。

【図 1 1】 目的地表示画面のスクロール時の様子を説明するための図である。

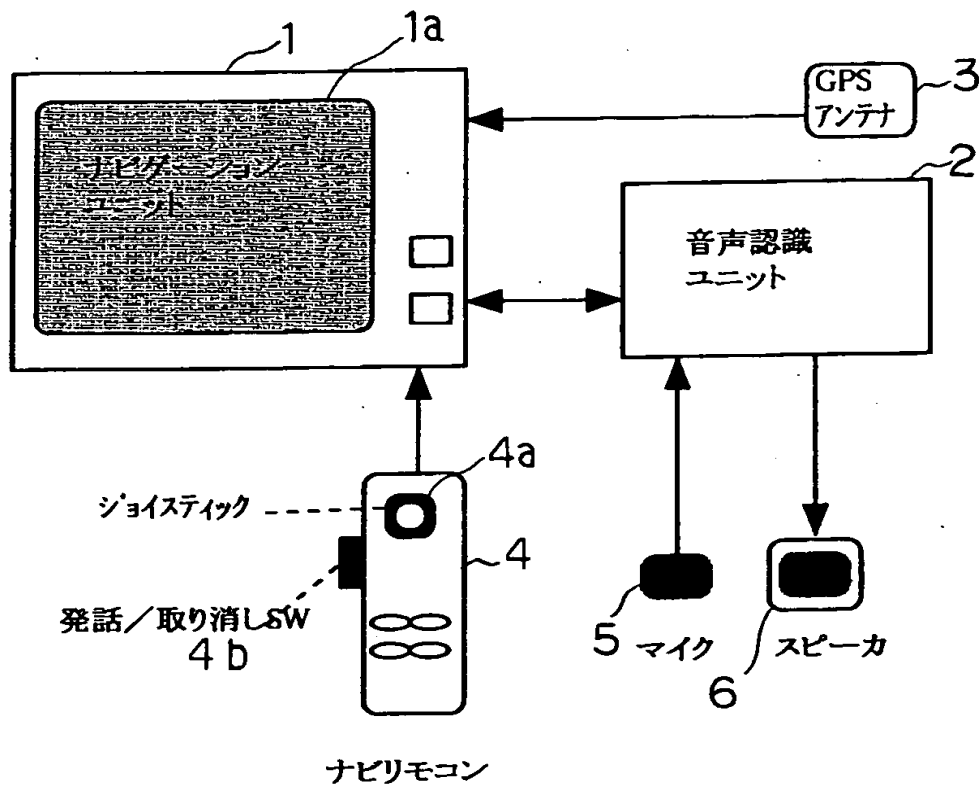
【図 1 2】 第 8 の実施の形態の認識対象語の重み付け方法を示す図である。

【符号の説明】

- 1 ナビゲーションユニット
 - 1 a モニター
- 2 音声認識ユニット
 - 2 a CPU
 - 2 b メモリ
 - 2 c 信号処理装置
 - 2 d A/Dコンバーター
 - 2 e D/Aコンバーター
 - 2 f 出力アンプ
 - 2 g 入力デバイス
 - 2 h 外部記憶装置
- 3 GPSアンテナ
- 4 ナビリモコン
 - 4 a ジョイスティック
 - 4 b 発話／取り消しスイッチ
- 5 マイク
- 6 スピーカー

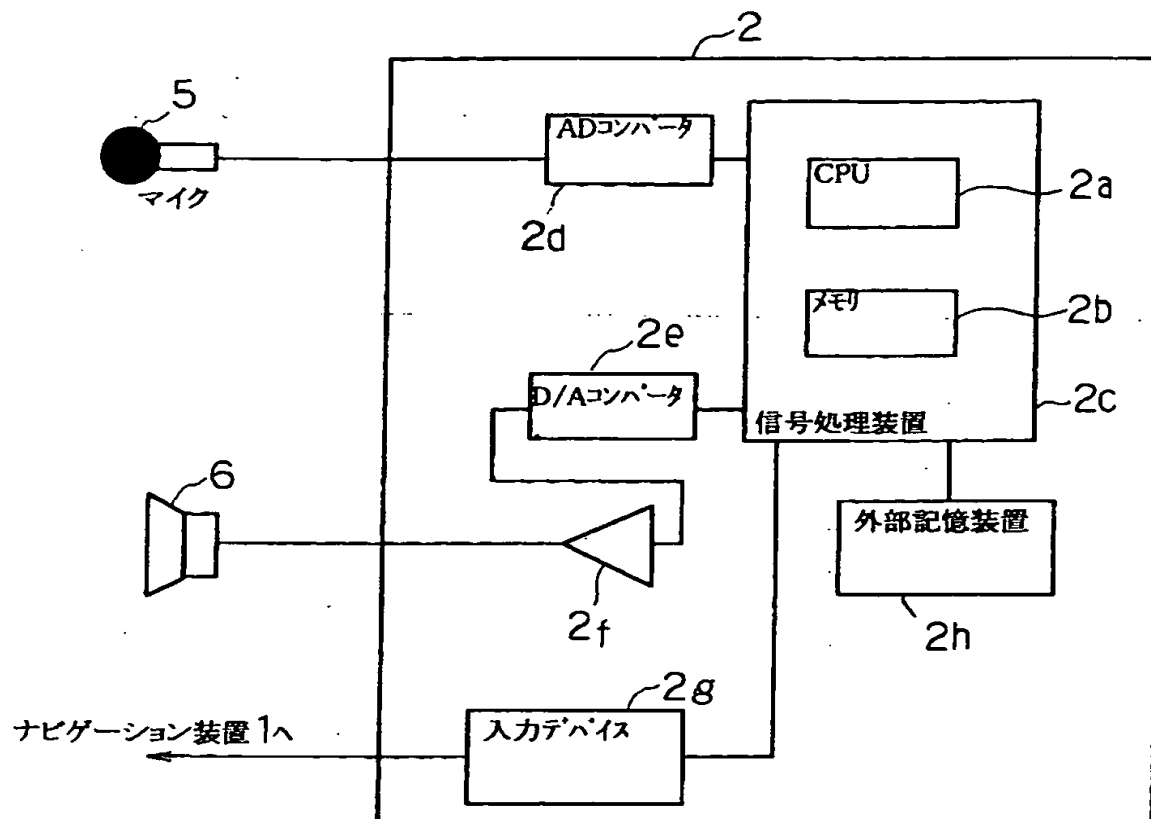
【書類名】 図面

【図 1】



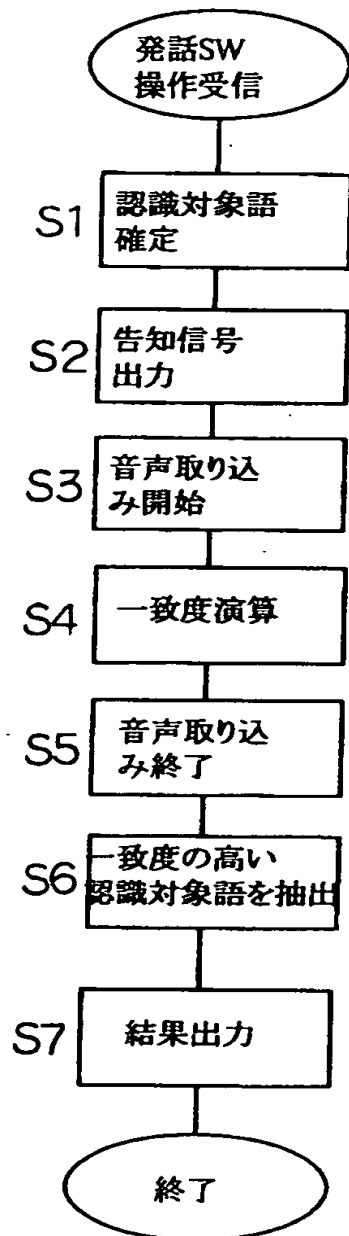
【図 1】

【図 2】



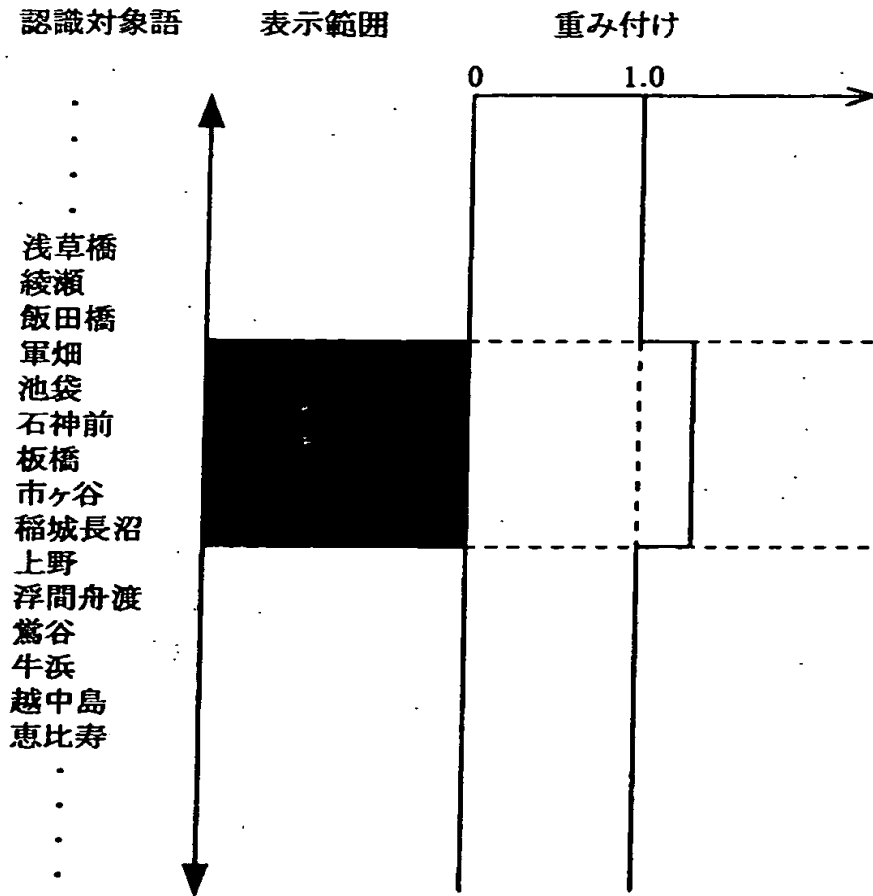
【図 2】

【図 3】



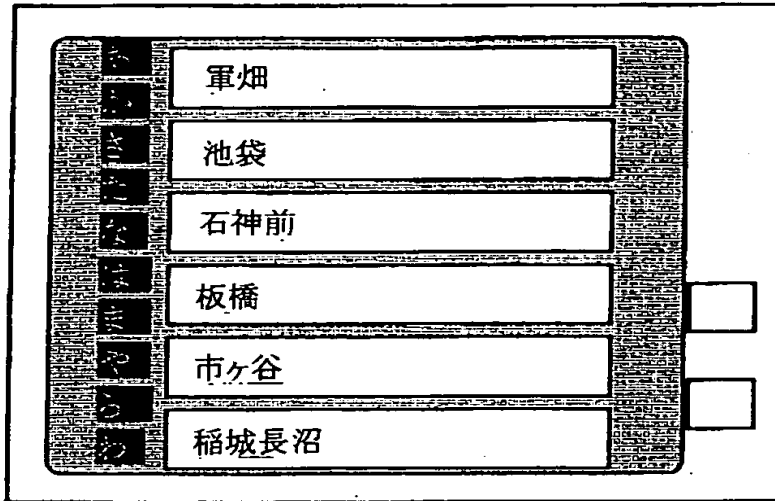
【図 3】

【図 4】



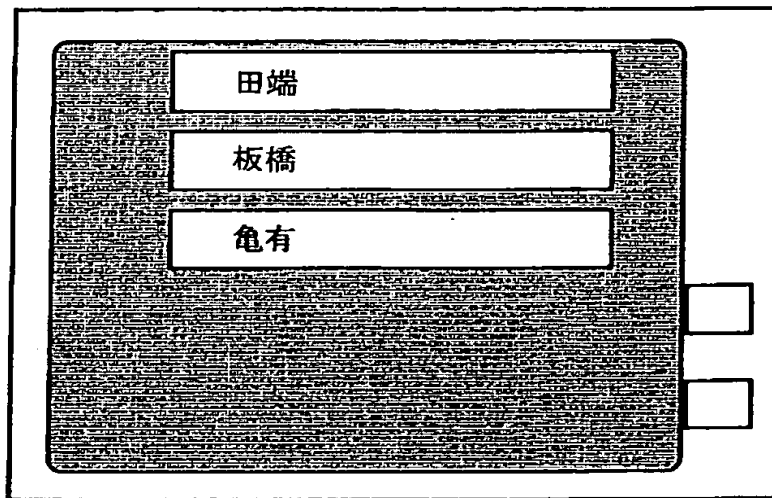
【図 4】

【図5】



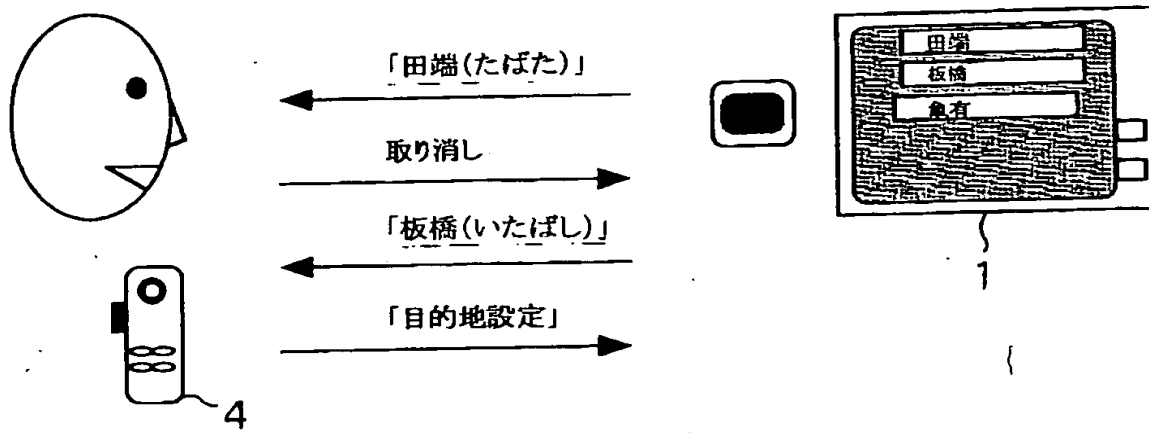
【図5】

【図6】



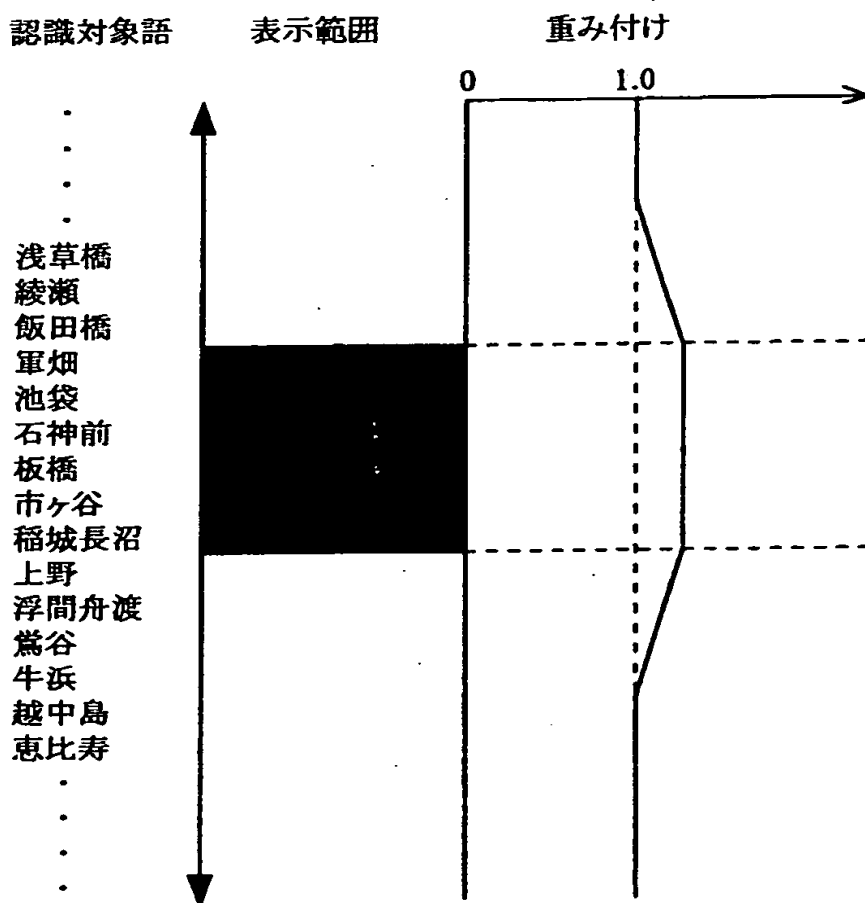
【図6】

【図 7】



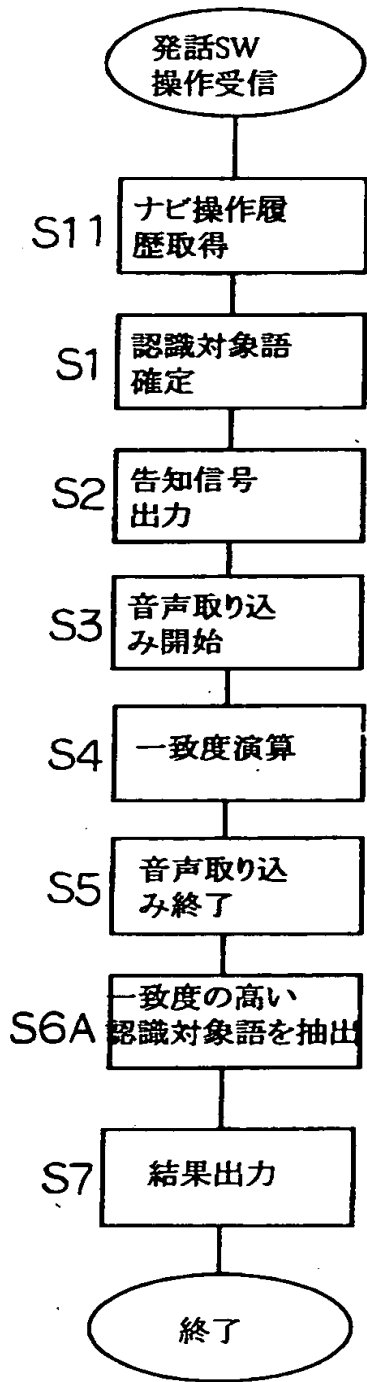
【図 7】

【図 8】



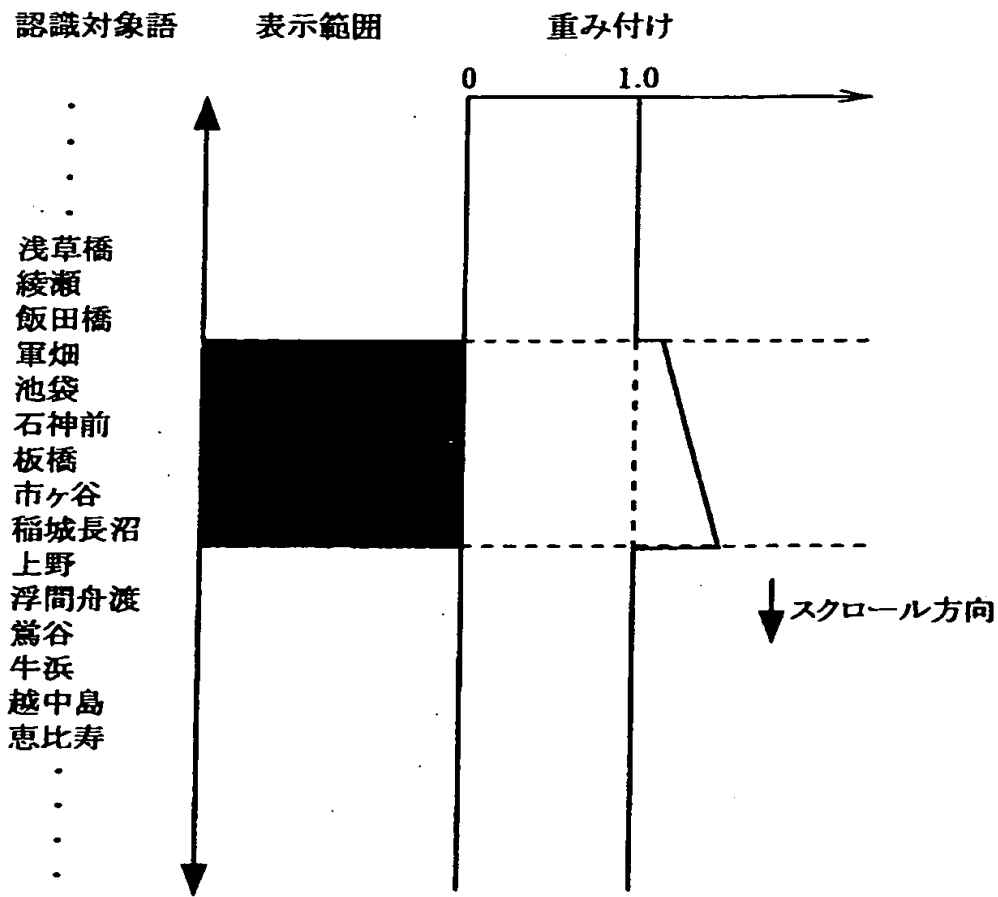
【図 8】

【図 9】



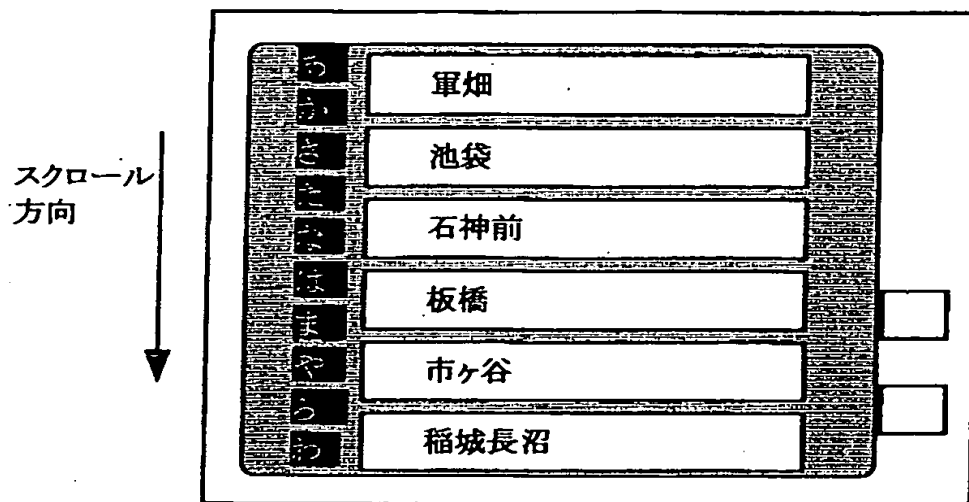
【図 9】

【図 10】



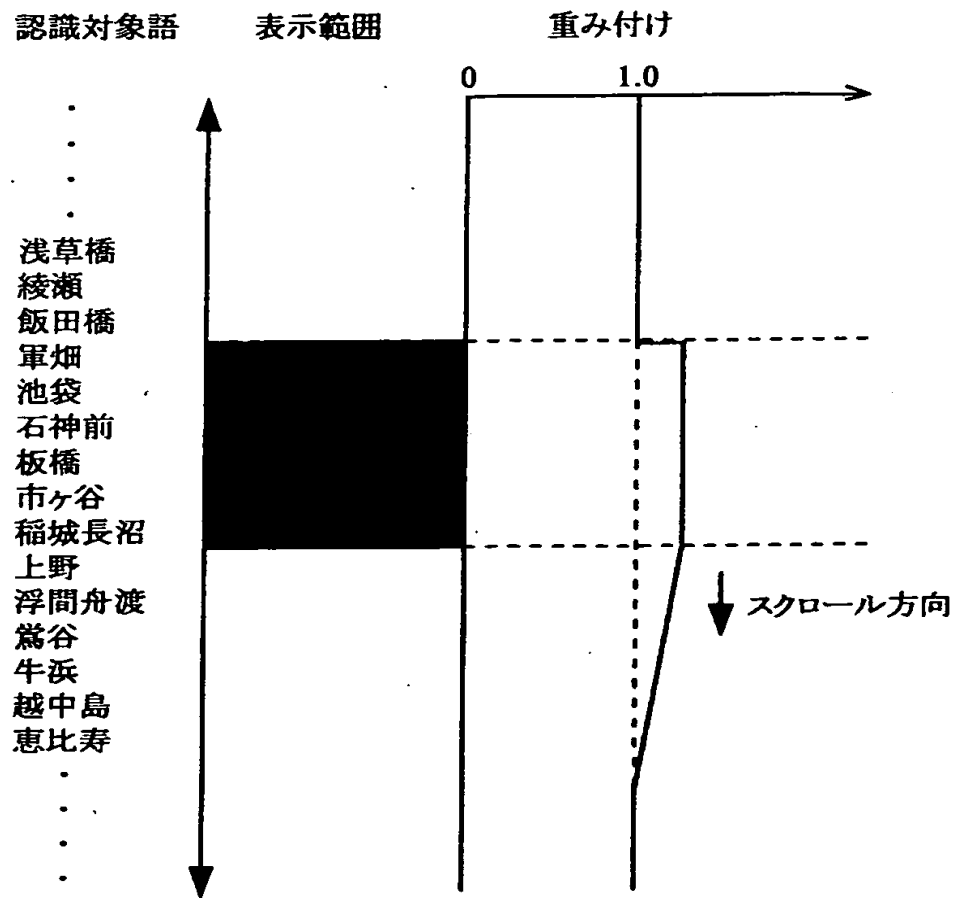
【図 10】

【図 11】



【図 11】

【図 12】



【図 12】

【書類名】 要約書

【要約】

【課題】 モニターに表示されている認識対象語に対する認識率を向上させる

【解決手段】 音声を入力する音声入力手段 5 と、複数の認識対象語を記憶する記憶手段と、記憶手段に記憶されている複数の認識対象語の内の、所定個数の認識対象語を表示する表示手段 1 a と、表示手段 1 a に表示される認識対象語に対して表示手段 1 a に表示されない認識対象語よりも高い重み付けを行う重み付け手段 2 と、重み付け手段 2 による重み付け後の認識対象語と入力手段 5 により入力された音声との一致度を演算する演算手段 2 とを備え、一致度の演算結果に基づいて入力した音声を認識する。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願2001-077910
受付番号	50100388570
書類名	特許願
担当官	第八担当上席 0097
作成日	平成13年 3月21日

<認定情報・付加情報>

【提出日】	平成13年 3月19日
-------	-------------

出 願 人 履 歴 情 報

識別番号 [000003997]

1. 変更年月日	1990年 8月31日
[変更理由]	新規登録
住 所	神奈川県横浜市神奈川区宝町2番地
氏 名	日産自動車株式会社